SOME FEATURES OF THE STATISTICAL WORK OF THE BUREAU OF LABOR STATISTICS.*

By Royal Meeker, Ph.D., *Commissioner of the Bureau.*

I agree with much that Professors Willcox and Durand have said. I believe that it is desirable to secure comparability and continuity in our statistical output. In revising the statistical work of the Bureau of Labor Statistics I have kept these things in mind and, so far as possible, have constructed the new series of price and wage indexes so as to admit of comparisons with the old indexes. The new wholesale price index number will be calculated back to 1890 so that the old index and the new may be compared throughout the whole period since 1890. The retail price index, however, has been calculated back only to and including 1907, for two reasons: First, because of the enormous amount of labor involved and the insufficient force and funds of the Bureau, and, second, because since 1907 retail prices have been much more accurately reported by merchants than previously, greater care being exercised to obtain the actual sale prices on the 15th of each month of the same grade of each commodity sold by each store. The price quotations before 1907 and since that date are so different that comparisons cannot fairly be made. No good purpose can be served by calculating the relative prices and index numbers back of 1907. Such a continuous series of relatives and indexes would give merely a fictitious comparability and continuity to retail price statistics, the prices themselves being discontinuous and incomparable.

It was my desire to carry back through 1907 the new index numbers showing changes in hourly rates of wages, full time weekly earnings, and weekly hours of labor in different industries. In some industries, however, the number of establishments reporting in 1911 was greatly increased over the number in 1910, so that no fair comparison could be made going back of 1911. It is unfortunate that the relatives and index

* Paper presented at a joint meeting of the American Economic Association and the American Statistical Association, Princeton, N. J., December 30, 1914.

numbers cannot be carried back in every instance to 1890, but I feel that it is better to make no comparisons of wages and prices rather than to make erroneous comparisons by means of a fictitiously continuous series of relative prices, wages, earnings, and hours of labor.

I sincerely hope that the proposed joint committee of the Economic Association and the Statistical Association, to advise with the statistical bureaus of the government, will be appointed. Unnecessary duplication of statistical work should be eliminated, and the statistical methods used should be standardized and made uniform. Especially is standardization and uniformization of methods needed in the different commonwealths of the United States. For example, the accident statistics of one state cannot at the present time be compared with the accident statistics of any other state because the statistical methods are so utterly at variance.

The various statistical bureaus in the federal service are getting together as never before so as to avoid duplication of effort and to agree upon the right things to do and the right way to do them. I am trying to prevail upon the various state agencies to coöperate with the Bureau of Labor Statistics and with each other in the gathering of statistics of accidents, unemployment, retail prices, wages, and hours of labor.

It is slow and discouraging work, but great good will result, if we can agree upon what statistics should show and the best methods of showing what is wanted. The committee suggested would be of great service in establishing proper statistical standards, in calling attention to the enormous quantity of costly and useless statistical output, and in eliminating duplication of work and the confusing and costly publication of more or less contradictory statistical stuff by various federal and state agencies.

Most treatises on statistics deal exclusively with statistical methods. Nothing is said about the data to which correct statistical methods are to be applied. Now all experience shows that the principal source of error in statistical work lies in the original figures collected to represent wages, prices, hours of labor, etc. It is much more important that we give attention to the gathering and verification of the original data,

than that we devote our efforts to hair-splitting refinements in the methods of treating the facts gathered. With all the treatises on statistical methods, however, we have not learned the difference between right methods and wrong methods in our statistical analyses. In fact, the treatises themselves are generally very vague or perfectly noncommittal as to the applicability of a given statistical method to a given set of conditions. Some writers seem to think that it really doesn't much matter what statistical methods we employ, as approximately the same results are obtained by any and all methods. I do not hold this comfortable view. I believe there is a right way and a wrong way of doing things. I believe it is of first importance to get our facts right, but I am convinced that the most irrefragible facts are sometimes made to tell lies because they are treated by wrong statistical methods.

Long before I took charge of the Bureau of Labor Statistics, I had become very suspicious of the Bureau's index numbers, especially its retail price index. Some people here present will no doubt recall that I was wont to have fun with the Bureau's index numbers. I no longer have fun with them—they have fun with me. As soon as I took charge of the Bureau, before I had got settled in the saddle, I set about to revise and recalculate the index numbers published by the Bureau. Perhaps some of you will recall receiving more than a year ago a letter from the Commissioner of Labor Statistics—a S. O. S. call for help. I did not have the self-assurance to set about revising the index numbers without first taking counsel with those who were best qualified to advise me in this matter. Many of you who received this distress signal may remember carefully depositing it in your waste basket. At least I received no reply from a number of economists and statisticians whom I addressed. Most of you who did reply may remember dictating a perfectly perfunctory reply which revealed either that you did not know or care much about index numbers, or that you did not wish to prolong your correspondence with the Commissioner of Labor Statistics. I want here and now to acknowledge publicly the assistance given me by Professor W. C. Mitchell and Professor Irving Fisher. Had it not been for the sympathy, encouragement, and counsel of Professors

Mitchell and Fisher, I should not have had the courage to carry out the recasting of the Bureau's index numbers. I have changed the methods employed in constructing the index numbers and shifted the base period from 1890–99 to the last completed current year. It has been a most laborious and tedious undertaking, but the work is now nearly completed.

That the method used in constructing index numbers is not an inconsequential matter is shown by a brief examination of the following table:

PRICES OF POTATOES FOR MAY, JUNE, AND JULY, 1913.

1. Firm.	May.		June.			July.		
	2. Price.	3. Relative on May base.	4. Price.	5. Relative on May base.	6. Relative on June base.	7. Price.	8. Relative on June base.	9. Relative on May base.
804	$0.20	100	$0.40	200	100	$0.30	75	150
808	0.17	100	0.36	211	100	0.32	89	188
815	0.50	100	0.40	80	100	0.35	87½	70
817	0.20	100	0.20	100	100	0.30	150	150
821	0.20	100	0.40	200	100	0.35	87½	175
City aggregates	$1.27	500	$1.76	791	500	$1.62	489	733
City relatives	100	100	139	158	100	92	98	147

City relative for July on May base computed by averaging and multiplying relatives..... 155
City relative for July on May base computed by comparing aggregate actual prices..... 128

This table gives the prices of potatoes during the months of May, June, and July, 1913. These prices have already been published by the Bureau of Labor Statistics. The prices were reported to the Bureau by five identical firms in one city—all the identical firms reporting for that city for the three months, May, June, and July.

In the first column is given the firm number. In the second column are given the money prices per peck for potatoes as reported by these five firms. In the third column the money prices are reduced to percentages or relative prices, May being taken as the base. Of course, all relative prices for that month are represented by 100. In the fourth column are given the June prices per peck for potatoes reported by these same five firms. In the fifth column are given the relative prices in June on the May base, that is, the percentage of June

prices to May prices. You will notice that some of the prices reported to the Bureau look very peculiar. Firm 808 reports potatoes at 17 cents a peck in May. Firm 815 reports potatoes by the peck at 50 cents in May. Clearly here we are dealing with different economic commodities. No doubt, 17 cents is the price for old potatoes, while 50 cents is the price for new potatoes. Note the phenomenal changes in prices in June as compared with May. The first firm shows an increase of 100 per cent. in the price of potatoes; the second firm shows an increase of 111 per cent.; and the third firm shows a decrease of 20 per cent. Now I wish to emphasize that these prices are actual prices reported to the Bureau by all of the identical firms in one city that reported for the three months under consideration. We instruct our retail grocers to report new potatoes only when the sales of new potatoes make up more than 50 per cent. of their total sales of potatoes. In this case it looks as if some of the firms had not strictly followed instructions, but I have as yet discovered no way of going behind the "election returns." We must rely upon the honesty and intelligence of the firms reporting. Had the prices reported by Firm 815 been called to my attention early enough, I should have eliminated it altogether—at least the price reported for May.

Now let us consider the city relative price constructed by averaging the individual firm relatives for the month of June. The simple arithmetic average of the relatives in column 5 is 158, that is, according to the information given by this relative price, prices of potatoes in this city have risen 58 per cent. from May to June. By comparing the aggregate money prices we get quite a different result. The aggregate prices may be used in the construction of relative prices in this case, because the firms reporting are identical for the months of May and June. Of course, if a different number of firms had reported in May as compared with June, it would be necessary either to compare identical firms or to reduce the aggregate prices to average prices before anything like an accurate result could be obtained. The aggregate money prices in May are $1.27. The aggregate for June is $1.76. Reducing both these money prices to percentages of the May price by

dividing by $1.27, we obtain as the June relative price on the May base, 139. This is almost 20 points less than the June relative price obtained by averaging individual firm relatives—a difference which certainly is not negligible and which is altogether due to the difference in method of calculating the relative prices.

Whatever may be said of the excellences of a general relative price constructed by the method of averaging relative prices built up from different bases for the purpose of showing changes in the cost of living, a relative price built up from actual money prices shows much more accurately what we want to show, namely, changes in the cost of living—changes in the cost of the same quantity of a commodity or of an unvarying market basket.

Since 1907 the method followed by the Bureau in constructing relative retail prices and index numbers has been as follows: Identical firms were compared month by month, the theory being that it is inaccurate to compare changes in relative prices of five firms one month, eight firms the following month, and ten firms the next month. In this way a relative price for February would be constructed on the January base by comparing the identical firms reporting both for January and February. Then this February relative price on the January base was multiplied into the January relative price on the base chosen for all relative prices and index numbers published by the Bureau, namely, the period 1890–99. Next a March relative price was constructed on the February base by comparing the prices of identical firms reporting for these two months. This March relative in turn was multiplied into the February relative price constructed on the 1890–99 base. In this way only identical firms were brought into comparison month by month. Now this method of bringing into comparison only identical firms is a perfectly good and accurate method, if properly applied, but applying it in this way necessitated shifting the base of the old index number every month. A relative price or index number built up by the method of averaging relative prices constructed on different bases cannot be shifted without a percentage of error that can only be guessed at. Every time the old index was shifted in

the way described above error was injected into the result, and the error was perpetuated and probably cumulated month by month and year by year. I will refer to this source of error a little later on.

Applying the method of comparing identical firms month by month to the figures in the table before us, we obtain an average relative price of 98 for July on the June base, for the five identical firms reporting in both June and July. By comparing the aggregate money prices reported by the five identical firms, we get the July relative price, 92, on the June base.

In column 9 are given the July relative prices constructed directly on the May base. The average of these relative prices is 147. By the method of shifting from one base to another we get a very different result. Multiplying the average relative price for July on the June base (98) by the average relative price of June on the May base (158) we obtain what purports to be the July average relative price on the May base (155). The difference between these two averages of relative prices for July on the May base is 8 points. In the first case the July prices are compared directly with the May prices. In the second case the July relative price on the June base is multiplied into the June relative price on the May base. This gives some idea of the possible discrepancies which may arise from using the method of comparing identical firms month by month, even when there is no change in firms whatsoever.

When we compare the aggregate money prices reported by identical firms in June and July, we obtain a July relative price of 92 on the June base. Now this relative price can be *shifted without error* to the May base or any base desired. This may be best shown by the following simple arithmetical formulæ: $\frac{\$1.76 = \text{June aggregate price}}{\$1.27 = \text{May aggregate price}} = \text{the June relative price}$ on the May base. $\frac{\$1.62 = \text{July aggregate price}}{\$1.76 = \text{June aggregate price}} = \text{the July}$ relative price on the June base. Multiplying the July relative price on the June base by the June relative price on the May base, we have $\frac{\$1.62}{\$1.76} \times \frac{\$1.76}{\$1.27} = \frac{\$1.62}{\$1.27} = 127\frac{71}{127}$, the July relative price on the May base, which is exactly the same result

as would be obtained by dividing the July aggregate in the first instance by the May aggregate. Even with changing firms and commodities varying in quality, the relative prices calculated by comparing actual prices may be shifted to any desired base by the method illustrated above with the closest possible approach to absolute accuracy. Shiftability is an indispensable quality in a relative price which must be built up by computing the relative for each month with the preceding month as the base and then shifting the resulting relative price to the selected base period by multiplying through by the relative price for the preceding month computed on the selected base period. If there were no other reason for changing the method of computation, this alone would seem to make the proposed change in method imperative.

The July relative price of potatoes on the May base computed by the old method employed by the Bureau is 155. The relative price of potatoes for the same month on the same base computed by the new method is 128. The difference is 27 points—a difference so great as to shake one's faith in relative prices and index numbers, if we had nothing to indicate to us whether the relative 155 was better or worse than the relative 128. In fact, however, a relative computed from actual money prices does reflect as accurately as possible the percentage changes in the cost of a given commodity. The relative 128 is, therefore, more trustworthy and exact than the relative 155.

In the same way a weighted index number of the family food budget, constructed by the use of actual money prices weighted according to the quantities of each commodity entering into consumption, is much more accurate and trustworthy than either an unweighted or a weighted index number constructed by the old method of averaging averages of relative prices to the fourth and fifth degree.

The advantage of constructing relative prices and index numbers which can be shifted to any base desired has still another important aspect. People are curious to know the percentage of price change from 1912 to 1913 or from 1907 to 1913, or for some other recent period of time. Few are interested to know by how large a per cent. the prices of 1913

exceed the prices of a period as remote as 1890–99. It is impossible by means of the old series of relative prices and index numbers to calculate accurately the percentage change in prices from 1912 to 1913. For example, from the Bureau's Bulletin 140, p. 16, we learn that the relative prices of round steak were 174.3 and 199.5, respectively for 1912 and 1913. Nobody can from these figures calculate the percentage of change in the prices of round steak from 1912 to 1913. The severest critics of the Bureau's price statistics almost invariably calculate the percentage of change by the short and simple process of subtraction, contenting themselves with the misinformation that the price of round steak rose 25.2 per cent. from 1912 to 1913. A more "scientific" method employed is to divide both relative prices through by the 1912 relative, 174.3, thereby going through the motions of shifting the base period to 1912, and obtaining 100 and 114.5 as the relative prices of round steak for 1912 and 1913, respectively, computed on the 1912 price as the base. The Bureau has resorted to this method in previous bulletins, to construct tables purporting to show the percentage changes in prices from year to year. This method of procedure is mathematically unsound and the result is vitiated by an amount of error that can be ascertained only by digging up the original price data and reconstructing the relative prices anew on the 1912 base. That the possible error is no negligible quantity is demonstrated by a brief consideration of the table below:

PRICES OF POTATOES FOR JUNE AND MAY, 1913.

Firm.	June.		May.	
	Price.	Relative.	Price.	Relative.
804	$0.40	100	$0.20	50
808	0.36	100	0.17	47
815	0.40	100	0.50	125
817	0.20	100	0.20	100
821	0.40	100	0.20	50
Aggregate	$1.76	500	$1.27	372
Relative price	100	100	72	74
Relative price obtained by shifting (100÷158)	.....	.....	.....	63

The relative price for June on the May base computed by averaging relative prices is 158. This quantity is supposed to

give the percentage relation that June prices bear to May prices. It is desired to find what is the percentage of May prices to June prices. Using the usual method of dividing through by the relative price (158) of the period to be used as the new base, we get the following relatives: June 100, May 63. When we compare the money prices and calculate the firm relative prices and average them we get 74, as shown above.

The relative price computed from the original price quotations is more than 17 per cent greater than the relative price obtained by shifting the base in the manner described above.

It must not be forgotten that the figures used are actual prices returned to the Bureau by all the identical dealers reporting from *one* city. This is not a case cooked up for the purpose of showing a theoretical possibility that contains no element of probability. I chose potatoes deliberately because their prices behaved so oddly at just this period when new potatoes are coming in and old potatoes are going out. The example given is extreme, but it is by no means unusual, and such capricious fluctuations are repeated every year for potatoes, and to a lesser extent for eggs and some other commodities that are subject to rather violent price changes. No doubt, more startling examples could have been found by a very little search. These examples are cited to show typical price changes in a commodity that fluctuates capriciously in price, not to exhibit the most extreme cases of such capriciousness.

The relative price computed from aggregate actual prices can be shifted at will to any base without error. This is evident when we consider the nature of such a relative price. The June relative price computed on the May price as 100 is $\frac{\$1.76.}{1.27}$ Shifting this series to the June base by dividing by the June relative price gives the following: May relative price $\frac{\$1.27,}{1.76}$ June relative price 100. Individual commodity relative prices can thus be shifted to the base price of any period desired without error because the relative prices are simple ratios of actual aggregate prices. Dividing through by the relative price of any year or period merely has the effect of

substituting the aggregate actual price for the base period as divisor in the formula for computing the relative price.

By the old method of computation, errors in price data were not only perpetuated but cumulated by means of the vicious method of averaging, to the fourth and fifth degree, averages of relatives calculated from different prices as bases and by the still more unallowable process of shifting every month the base of the relative prices, which could not be done without subjecting the relative prices to grave suspicion as a dependable means of representing accurately what was happening to prices. These inaccuracies, taken with the inflexibility of relative prices and indexes calculated by averaging relatives, made the changes in methods of calculation which have been carried out imperatively necessary.